REMARKS

In the application, Claims 1 and 15-28 and 35-41 are pending and rejected. The rejections raised in the Office Action of February 25, 2004 have been considered and, in response, Claims 1, 21 and 35 have been amended. Applicant wishes to thank the Examiner for withdrawing the final rejection of the prior Office Action. It is submitted that the claims now presented are in a condition for allowance and the Examiner is requested to reconsider the claims and issue a notice of allowance.

Rejections under 35 U.S.C. §102

The Examiner rejects Claims 1, 15, 16, 21-24, 28 and 35-37 under 35 U.S.C. §102(b) as being clearly anticipated by Eckman et al.

Applicant respectfully submits that Eckman et al. teach a method for data mining EST (expressed sequence tag) sequence data, where the data consists of sequence listings of cDNA fragments along with associated sequence annotation and other genomic data. The searching options described by the authors are EST accession number, text search for gene name or index class identity. (See Fig. 3 and text at page 8, column 2 through page 10, column 1.). The EST accession number, if not previously known, is obtained by conducting a sequence similarity search using BLAST or a similar pairwise comparison to a comprehensive nucleotide database. Using the available combination of search functions, one can only conduct a search based on gene sequence or name. This limited capability arises at least in part from the fact that ESTs themselves are not gene expression data as described in the present application, rather, they are a means for obtaining gene expression data. More specifically, as applied to the present invention, the DNA microarrays used to generate gene expression data can be made using cDNAs or oligonucleotide probes linked to the surface of an appropriate substrate. The cDNAs or oligonucleotide probes may be derived from either full length genes or ESTs. To determine differences in gene expression, labeled cRNAs or cDNAs are then hybridized to the DNA-carrying microarrays. The expression of tens of thousands of genes can be analyzed on a single microarray by detecting the hybridization intensity. *This* is the gene expression data that is stored in the gene expression database of the present invention, not merely the sequence listing.

The Examiner points to a section of Eckman et al. in which the writers identified 12,000 distinct genes by focusing on differential tissue expression. There is no evidence from Eckman et al. that the disclosed database system contains gene expression data (i.e. transcript abundance). Rather, it appears that information obtained from the differential tissue expression study was used solely to identify unique sequence information. The purpose of the Eckman et al. system is to "produce a gene index to the human genome [the Merck Gene Index (MGI)], a non-redundant set of clones and sequences, each representing a distinct gene" (page 2, column 2). The database system structure figure (Figure 1, page 4) and the data descriptions by system databases (page 4, columns 1 and 2) do not suggest that the system contains information about transcript abundance. For example, the Merck Gene Index database is described as containing the EST's accession number, an integer code that reflects the results of the index quality analysis, the starting and ending position of index-quality sequence, the length of the index quality sequence, and the FastA bestscore (perfect match score) (page 4, column 1). The LENS database is described as containing a clone's IMAGE id, array address, insert size and GDB id, the name and GDB id of the library from which the clone was derived, the GenBank assession number, WashU id, and orientation of EST sequences (page 4, column 2). The dbEST database is described as storing the results of pairwise BLAST comparisons (page 4, column 2). As described, the databases do not contain gene expression data.

Furthermore, in Eckman et al.the procedure by which differential tissue expression was measured was 3' sequence clustering, which is described by the following excerpt from a portion of the website referenced in the article ((Matsubara, 1995) – http://bodymap.ims.u-tokyo.ac.jp/doc/detailed_explanation.html):

> To study expression profile of genes in a given cell or tissue with the currently available technology, a 3'-directed cDNA library is prepared, then randomly selected clones are sequenced in a large scale. The 3' region of mRNA, just upstream of the poly A stretch in mRNA, can be quantitatively converted to cDNA, and therefore, the 3'-directed library is a faithful representative of the molar composition of the mRNA. Since the sequences at the 3'-region are unique, sequencing data of about 150-300 nucleotides are sufficient to characterize the gene. This short sequence is called the "gene signature". . . . . The abundance of the mRNA can be measured by counting the number of appearances of the same "signature". With the help of automated sequencers, thousands of gene signatures, with their frequency profiles, can be collected with the mRNA population of a given cell or tissue. The resulting data is referred to as the "gene expression profile" of the tissue.

In contrast, gene expression data generated using microarrays consists of hybridization signal level, i.e., fluorescence intensity, as measured using a detector. These results can be interpreted rapidly and, due to the number of probes on each array, are capable of generating very different and far greater volumes of distinct data than the above-described sequence-counting approach. As a result, the data handling requirements are significantly different and sequence searching ability alone is inadequate to effectively mine the data. Accordingly, the fact that Applicant's invention is drawn to a method and system for managing gene expression data derived from DNA microarrays raises different problems requiring different solutions. Applicant respectfully submits that Eckman et al. do not teach a database comprising gene expression data obtained from DNA microarrays, but only a database of sequence listings and, as a result, cannot anticipate the elements of Applicant's invention.

Additionally, Eckman et al. do not teach a system that provides multiple databases for separately storing different categories of data as taught by Applicant. While Eckman et al. do describe multiple databases, there are not three distinct databases taught for separately storing each of gene expression data, gene annotation data and sample data as claimed by Applicant. As previously pointed out, Eckman et al. do not teach a database containing gene expression data. Eckman et al. teach a three component gene annotation database which contains a minimal amount of sample information in the form of library information in the LENS database (page 4, columns 1 and 2). The lack of distinction between the different attribute categories for gene expression data, sample data and gene annotation data severely limits the scope and flexibility of a search using the Eckman et al. system compared with Applicant's system and method. As such, the three separate databases are an important aspect of Applicant's invention which is not disclosed by Eckman et al..

For the foregoing reasons, the Eckman, et al. reference does not teach each and every element and, therefore, cannot anticipate Applicant's invention as now claimed. Accordingly, Applicant respectfully requests that the Examiner withdraw the rejection under §102(b).

Rejections under 35 U.S.C. §103

The Examiner rejects Claims 1, 15-28 and 35-41 under 35 U.S.C. §103(a) as being unpatentable over Eckman et al. taken with Schena et al. (1996).

Applicant respectfully submits that the data in the databases disclosed by Eckman et al. are limited to sequence-based data and gene/EST identifier data. As such, there is no motivation to separate the data into three distinct databases as taught by Applicant to enable searching beyond text or sequence-based data. In microarray-based expression analysis, a single microarray can produce tens of thousands of expression data points. Given the great quantity of expression data and the fact that it is distinct from sequence data and gene/EST identifier data, the method disclosed by Eckman et al. is inadequate to provide the database management and mining capability provided by Applicant's invention.

The Examiner states that one of ordinary skill in the art would have been motivated by the improvement disclosed by Schena et al. to utilize the vast amounts of EST data disclosed by Eckman et al. for sequence analysis and gene discovery. While the Schena et al. reference discloses methods for microarray-based expression monitoring, and, thus, recognizes the differences between EST sequence-based expression data and DNA microarray hybridization-based expression data, it fails to provide any teaching or suggestion relative to the handling and analysis of such data. At page 10615, column 1 (Computer Graphics and Informatics), Schena et al. describe the use of commercially-available image analysis software for analyzing differential expression representations. Other than detection and image analysis, all that is describe under this section is sequence searching, which is taught by Eckman et al. Schena et al. do not address any data management or mining techniques because their research involved discrete tests of high-throughput gene expression monitoring. Their goal was not to create a database that was suitable for data mining. As such, they provide no teaching or motivation for defining three distinct databases for separately storing gene expression data, sample data and gene annotation data. Neither Eckman et al. nor Schena et al. recognize any advantage to the ability to conduct a search based on more than sequence data or descriptive text. In contrast, Applicant's database configuration permits a wide variety of search approaches based on any one of gene expression results, sample data or gene annotation, or some combination thereof. Consequently, a more versatile and potentially useful search can be performed using Applicant's invention.

It is respectfully submitted that the database management method now claimed is not obvious from Eckman et al. alone or in combination with Schena et al. The database described by Eckman et al. is a research or reference tool which requires queries based on sequence data

and information directly related to that sequence data. Schena et al. describe no database -- only a kind of data. Because of the separation of data types into three distinct databases and the common interface that links them, Applicant's invention is a powerful discovery tool that allows one to take numerous varied approaches to data mining for virtually unlimited exploration of DNA microarray-based gene expression data and related data. Accordingly, Applicant respectfully submits that the invention as now claimed is patentably distinct over the cited combination and requests that the Examiner withdraw the rejection under §103(a).

Conclusion

In view of the foregoing amendments and remarks, Applicant submit that all bases for rejection have been addressed and overcome such that the amended claims are allowable over the prior art. Accordingly, Applicant respectfully requests that the Examiner withdraw all rejections set forth in the Office Action and issue a notice of allowance for all claims now in the application.

Should the Examiner believe that prosecution of this application might be expedited by further discussion of the issues, he is invited to telephone the undersigned attorney for Applicant at the telephone number indicated below.

Respectfully submitted,

Dated: 25 Aug. 2004

By: 

Eleanor M. Musick
Attorney for Applicant
Registration No. 35,623

Procopio, Cory, Hargreaves & Savitch LLP
530 B Street
Suite 2100
San Diego, California 92101
Telephone: (760) 931-9700
Facsimile: (760) 931-1155
E-mail: emm@procopio.com

Docket No. 4010US (111944/000015)